Influential Factors on Incident Management:
Lessons Learned from a Large Sample of Products in Operation

João Caldeira, Fernando Brito e Abreu

Faculty of Sciences and Technology
Universidade Nova de Lisboa
Lisboa, Portugal
j.caldeira@fct.unl.pt, fba@di.fct.unl.pt

Abstract.
Understanding causal relationships on incident management can help software development organizations in finding the adequate level of resourcing, as well as improving the quality of services they provide to their end-users and/or customers. This paper presents an empirical study conducted upon a sample of incident reports recorded during the operation of several hundred commercial software products, over a period of three years, on six countries in Europe and Latin America. The underlying research questions refer to the validation of which are the influencing factors affecting the incidents management lifecycle. Nonparametric analysis of variance procedures are used for testing hypotheses.

Keywords: Software development, Empirical Software Engineering, Software Quality, ITIL, Incident Management, Problem Management, Release and Deployment Management, Services Science

1 Introduction

> *"If you want the present to be different from the past, study the past."*
> Baruch de Spinoza (1632-1677)

1.1 Motivation

Organizations with in-house software development strive for finding the right number of resources (with the right skills) and adequate budgets. A good way to optimize those figures is avoiding expenditures on overhead activities, such as excessive customer support. This can be achieved by identifying incident's root causes and use that knowledge to improve the software evolution process.

Software development and software quality improvement have been strong topics for discussion in the last decades [1, 2]. Software Engineering has always been concerned with theories and best practices to develop software for large-scale usage. However, most times those theories are not validated in real-life environments [3]. Several factors were identified that explain this lack of experimental validation [4].

In real-life operation environments end-users/customers face software faults, lack of functionalities and sometimes just lack of training. These incidents should be somehow reported. According to the ITIL[1] framework [5], in an organization with a Service Management approach [6-10], this problem is addressed by two specific processes: *Incident Management* [6], which deals with the restoration of the service to the end-user within the Service Level Agreements [7, 10] (if they exist), and *Problem Management* [6] which aims at finding the underlying cause of reported incidents.

When an organization implements these ITIL processes, then it will address all kind of incidents (software, hardware, documentation, services, etc) raised by the end-users/customers. In this paper we are only concerned about software-related incidents.

The incidents database is an important asset for software engineering teams. Learning from past experience in service management, allows shifting from a reactive approach to a more proactive one. The latter is referred in the Software Maintenance chapter of the SWEBOK[2] (see Table 1), although seldom brought to practice.

Table 1. Software maintenance categories (source: SWEBOK [11])

	Correction	Enhancement
Proactive	Preventive	Perfective
Reactive	Corrective	Adaptive

This paper presents a statistical-based analysis of software related incidents resulting from the operation of several hundred commercial software products, from 2005 to 2007. The incidents were reported by customers of a large independent software vendor. Although that vendor operates worldwide, we were only able to have access to data from six countries in Europe and Latin America. Further details regarding the products and their users cannot be provided here due to a non-disclosure agreement.

The main goal of this paper is shedding some light on the influential factors that affect incidents lifecycle from creation to its closure, namely the schedule of its phases. Understanding this lifecycle can help software development organizations in allocating adequate resources (people and budget), increasing the quality of services they provide and finally improving their image in the marketplace.

The work presented herein is on the crossroads of Empirical Software Engineering and of the emerging area of Services Science [12, 13]. It is organized as follows: section 2 presents a survey of related work; section 3 contains the empirical study; finally, section 4 presents the conclusions, the threats to validity and future work.

---

[1] - The IT Infrastructure Library (ITIL) is the *de facto* standard for IT service management, an initiative launched in the late 80´s by the UK Office of Government Commerce's (OGC). The ITIL framework is as a generic reference model proposing a set of concepts and good practices for managing information technology, infrastructure, development and operations.

[2] - The Guide to the SoftWare Engineering BOdy of Knowledge (SWEBOK) is an IEEE CS initiative aiming to: (i) promote a consistent view of software engineering worldwide, (ii) clarify the place and set the boundary of software engineering with respect to other disciplines such as computer science, project management, computer engineering, and mathematics, (iii) characterize the contents of the software engineering discipline, (iv) provide a topical access to the Software Engineering Body of Knowledge and (v) provide a foundation for curriculum development and individual certification and licensing material.

2 Related Work

To support our research, we have tried to find related work in the area of empirical software engineering within the ITIL scope. Having searched several digital libraries such as the ones of ACM, IEEE, Springer or Elsevier, we were able to find only a few papers about incident management. Even scarcer were those referencing real-life empirical studies on software incidents and how they that can help improving the software engineering process. This section presents a categorized overview of the published works that we found to be closer-related to our work presented hereafter.

2.1 Categorization process

ITIL is concerned about three basic aspects in IT Service Management (ITSM): **technology**, **people** and **processes**. The technology aspect refers to all the technical components (typically hardware and software) involved when dealing with IT services. The people aspect addresses the way persons are organized and the way they should behave when involved in a certain process. Finally, the process aspect relates to how activities are linked together in order to deliver value to a specific business area. We categorized the related work according to the extent it has approached those ITSM aspects. For classifying each of the aspects, we use the following ordinal scale:

Absent / **Fuzzy view**	○	The topic is not addressed or addressed in a fuzzy way
Partly / **Isolated view**	◔	The topic is addressed insufficiently, not explicit or lacking context
Largely / **Contextualized view**	◑	The topic is addressed explicitly and context is provided, although not exhaustively
Fully / **Holistic view**	●	The topic is addressed exhaustively, sustained with evidence and with adequate rationale being provided

Besides that categorization, we provide, for each work, its main goal (as we perceived it), a commented abstract and, finally, we comment about the relation each work has with ours. Notice that we have kept the capitalized denomination of ITIL processes (e.g Incident, Problem or Configuration Management).

2.2 Review of related work

<u>Barash et al. (2007)</u> [14]

Technology	People	Processes
◔	◑	◑

Goal – Managing service incidents and improving an IT support organization
Comments – This work has a clear link with ITIL. The main topics addressed are Incident and Problem Management and the improvement an organization can achieve in their support activities by analyzing incident metrics. The authors suggest ways to improve staff allocation, shift rotation, working hours and the escalation of incidents.

We could not find, in this work, a clear link between Incident or Problem Management processes with the software development process and how they can help

each other in improving the quality of the service to the end-users. We also could not find a direct relationship to any other ITIL processes beyond the two referred ones. Nevertheless, we should not forget that if we improve the performance of the IT support organization, we are indirectly improving the performance of all other areas.

Relation with our work – This work is related with our own since it also addresses the management of incidents (herein we only address software incidents), and it tries to improve an IT Support Organization.

Sjoberg et al. (2005) [3]

Technology	People	Processes
○	◑	◑

Goal – A survey of controlled experiments in Software Engineering

Comments – In this work there is a detailed classification about the areas where those software experiments were conducted. It is interesting to realize that among the group of areas with fewer experiments, we find Strategy, Alignment, IT impact. These are within the most important issues addressed by ITIL and Service Management. One of the things that first came to our eyes is the fact that there is no category named "Service". We can assume that within all experiments done, none was made having the "Service" in mind. This is even more important since nowadays services are heavily dependent on software, and, on the other hand, the use of software can be seen as a service on its own. Overall, this work is a quantitative summary of controlled experiments. While the people and the processes aspects are briefly addressed, the technology aspect is only slightly covered. Indeed, few environment descriptions are provided on the technical conditions on which the experiments took place.

Although this survey was performed around three years ago, we have not found evidence, since then, contradicting the obvious need of more experiments relating software, services and their management processes.

Relation with our work – We expected that other studies like the one performed in our paper would be reported in this survey. While on the methodology side this is true, since many of the reported experiments use empirical data and statistical analysis, the same cannot be said regarding the context (incident management).

Niessink and Vliet (2000) [15]

Technology	People	Processes
◑	◑	◑

Goal – Software maintenance and software development from a service perspective

Comments – The authors clearly identify differences between services and products and how these differences affect the way end-users or customers assess their quality. One of the more relevant aspects of this work is the focus put on the need for defining Service Level Agreements (SLA), Service Catalogs and the importance of good Incident and Problem Management processes within an organization. These three aspects and the positive impact they can have in organizations that implement them are highlighted and understood, but not exhaustively explained. This would be addressed by detailing and giving examples on the implementation of the above aspects. In brief, the important topics are there, but not enough detail is provided.

Relation with our work – The relation lies on the ITIL focus. This is not an empirical study, but it covers all the important aspects of Service Management.

Jansen and Brinkkemper (2006) [16]

Technology	People	Processes
◑	○	◐
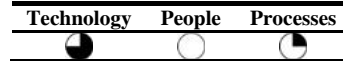
Goal – Study of the release, delivery and deployment of software

Comments – This is a very interesting paper about the software update process and how it can help software vendors and end-users/customers in the software deployment process. The approach taken fits in the realm of the Asset and Configuration Management, Release and Deployment Management ITIL processes. Notice that the deployment phase, which is focused in this paper, is precisely the one when most incidents are usually reported. This is due to the fact that IT systems and platforms are becoming increasingly more heterogeneous and complex and also because quality management systems (in general) and SLA verification (in particular) imply the recording of incidents originated by the operation.

Relation with our work – This work focuses on the technology used to improve the software deployment process, but does not cover any empirical study or data analysis. It is related to our work because it touches another key process in ITIL.

Mohagheghi and Conradi (2007) [17]

Technology	People	Processes
◑	○	◐
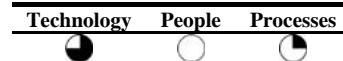
Goal – Quality, productivity and economic benefits of software reuse

Comments – This work is about software reuse and its benefits. Based on previous studies, the authors state that component reuse is related with software with fewer defects. The latter are identified by means of failures in operation and are the origin of reported incidents. The end-user perspective is not covered in this paper, and this is vital for a Service Management approach. Some references are made to software changes, software deployment and even infrastructure resources required for software execution. These are somehow implicit references to ITIL Change Management, Release and Deployment Management and Capacity Management processes.

Relation with our work – This work shares our objective of achieving a tangible and positive impact on the software development process by adopting ITIL-like best practices. This has strengthened our conviction that the impact of incident management on the software development lifecycle deserves further analysis.

2.3   Review summary

It is widely accepted that we lack experimentation in Software Engineering in general. This phenomenon is even more acute on what concerns experimentation related with incidents and services. As Spinoza observed more than 300 years ago, we need to understand how services were provided in the past to improve their quality in the future. Even if the related work is scarce, we should look at it collectively to try drawing some picture of the current state-of-the-art. For that purpose, a summary of the categorized related work is presented in Table 2.

Table 2. Summary of related work

Proposal	Technology	People	Processes	Relation
Barash et al. (2007)	◑	●	●	High
Sjoberg et al. (2005)	○	◑	◑	Medium
Niessink and Vliet (2000)	●	●	●	High
Jansen and Brinkkemper (2006)	●	○	◑	Low
Mohagheghi and Conradi (2007)	●	○	◑	Low

Out of the three aspects, the one that deserves the least attention is clearly "people", while the "technology" and "process" aspects have somehow equivalent emphases. We believe that this difference is due to the fact that researchers working in this area have mostly an Engineering background. Understanding people and their motivations requires Social Sciences skills.

However, the most relevant conclusion we reached while performing this unambitious state-of-the-art survey, is that the empirical study of incident management has not yet been adequately addressed in the scientific literature. We believe this situation is due to the fact that real-life samples contain sensitive data to companies and so are usually unavailable to researchers.

3 The empirical study

3.1 Process and instrumentation

Fig. 1. Empirical study workflow

Our empirical process consisted on the four steps represented in Fig. 1. We collected the data on the first days of January 2008, using an incident management system client interface. This tool allowed to export incidents data into a CSV (Comma Separated Values) file that could be loaded into a spreadsheet (MS Excel). Next, we filtered out a very small percentage of cases that had erroneous data (e.g. invalid dates). Then, we computed several variables from existing data, namely by calculating differences between pairs of dates.. The resulting dataset was then loaded into the SPSS statistical analysis tool, where the statistical analysis took place,

3.2 The sample

The subjects of our empirical study are around 23 thousand incidents, reported by end-users/customers, occurred during the operation of around 700 software products[3].

---

[3] - When a given product is available on different platforms, this number considers those instances as distinct products. Some distinction is also due to different licensing schemes.

The incidents were recorded with a proprietary incident management system during a time span of three years (2005 to 2007) in around 1500 companies in 6 countries.

We considered three geographical zones, with two countries in each one. The zones are Latin America (LA), Southwestern Europe (SE) and Central Europe (CE). Notice that there are 4 languages spoken in the considered countries: English (EN), French (FR), Portuguese (PT) and Spanish (ES). More details are provided in Table 3.

Table 3. Countries with their zones and languages

Country	Zone	Language	# of Incidents	# of Customers	# of Software Products
England (UK)	CE	EN	7349	530	460
France (FR)	CE	FR	8237	554	444
Spain (ES)	SE	ES	4014	219	359
Argentina (AR)	LA	ES	535	66	88
Portugal (PT)	SE	PT	556	37	107
Brazil (BR)	LA	PT	2221	125	250
Total			**22912**	**1531**	

3.3 Descriptive variables

Table 4. Variables used in this empirical study, their scale types and description

Variable	Scale	Description
Product	Nominal	Name of the product causing the incident
Company	Nominal	Name of the company where the product is installed
Country	Nominal	Name of the country where the incident was originated
Zone	Nominal	Zone of the globe where the country lies
Language	Nominal	Language spoken in the country
Category	Nominal	Represents incident's root cause
		Valid values are: 3rd Party Solution, Customer Support, Customization, Documentation, Function, Installation, Internationalization, Compatibility, Licensing, Localization, Performance, RFI, Security Threat, Stability, Education, Uncategorized
Impact	Ordinal	Measures incident's business criticality
		Valid values are: 1-Critical, 2-High, 3-Medium, 4-Low
Priority	Ordinal	Measures incident's correction prioritization as seen by the support[4]
		Valid values are the same as for the impact
Status	Nominal	Current status of the incident in its life cycle
WeekOfCreation	Interval	Order of the week (in the year) when the incident occurred
		Valid values belong to the interval [1, 53]
WeekdayOfCreation	Interval	Order of the day (in the week) when the incident occurred.
		Valid values belong to the interval [1, 7]
TimeToRespond	Absolute	Elapsed time from incident creation until a support person has started to work on it
TimeToResolve	Absolute	Elapsed time from incident creation until a resolution is given to the end-user
TimeToConfirm	Absolute	Elapsed time since the resolution was given to the end-user until a confirmation is obtained that the incident is closed

---

[4] - According to ITIL, incidents priority should be calculated based upon urgency and impact. However, the incident management system used in this study does not yet support the concept of urgency. The priority is assigned directly by the incident recorder.

The variables used in this empirical study are self-described in Table 4. The choice on the characterization of the incidents (*Category, Impact* and *Priority*) is performed by the person who registers the incident (the end-user/customer or a support staff member). Incidents have a defined lifecycle. In this paper we will only consider closed incidents, since those are the only ones for which we know the values of all timing variables. Fig. 2 describes how the three timing variables are calculated, regarding specific milestones in the incidents' lifecycle.
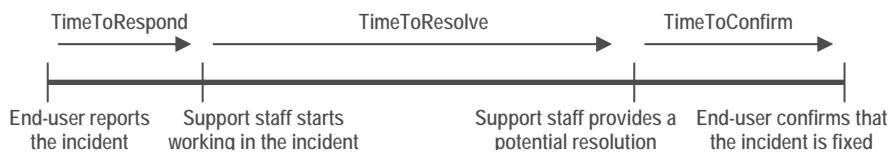

Fig. 2. Incidents' lifecycle timing variables

3.4 Research questions

To understand incident management we must find answers for these two questions:

    Q1: Which factors influence the lifecycle of incidents?

    Q2: Are there patterns in the occurrence of incidents?

Regarding Q1, the set of variables that best describe incidents lifecycle at a macroscopic level are *TimeToRespond, TimeToResolve* and *TimeToConfirm.* The answer to Q1 is important both to clients and service providers. For clients, particularly for large organizations operating in several countries, it will allow taking decisions in the formulation and negotiation of Service Level Agreements (SLAs). For service providers it will also help in finding the adequate level of staffing.

Regarding the possible factors influencing the incidents lifecycle, we can consider the following variables inscribed in Table 4: *Product, Company, Country, Zone, Language, Category, Impact* and *Priority.* We have selected the following research questions within the scope of this paper:

- **Has the impact of an incident an influence on its lifecycle?**

- **Has the priority of an incident an influence on its lifecycle?**

- **Has the originating country of an incident an influence on its lifecycle?**

- **Has the originating geographical zone of an incident an influence on its lifecycle?**

- **Has the language spoken in the country where the incident was reported, an influence on its lifecycle?**

- **Has the incident category an influence on its lifecycle?**

Regarding Q2, the occurrence of incidents can be measured by a simple counting or a weighted sum (e.g. taking the *Impact* or *Priority* as a weight) of incidents matching one of the possible values of the variable under consideration. For instance, if we were concerned with the identification of seasonal patterns, we can consider the day

within the week (*WeekdayOfCreation*) or the week within the year (*WeekOfCreation*) when the incidents were reported. Again, the answer to Q2 will bring benefits to client and service provider. Both will become aware of worst and best-case scenarios and thus take appropriate actions.

Due to the lack of space, we have just considered here a possible pattern, which is the distribution of critical incidents, the ones which give more headaches to all stakeholders. In this case, since the incidents were recorded using the same incident management system and supposedly using similar classification criteria, we would expect the proportion of critical incidents to be the same across countries. In other words, the corresponding research question is simply:

- **Is the distribution of critical incidents the same across countries?**

3.5   Hypotheses identification and testing

In this section we identify which are the statistical hypotheses that must be tested in order to answer the previously stated research questions. We then apply the adequate statistical tests and interpret their results. Research questions are prefixed by "**RQ**".

RQ: Has the impact of an incident an influence on its lifecycle?
In other words, we want to know if incidents with different assigned impacts differ in the corresponding lifecycle schedules (*TimeToRespond, TimeToResolve, TimeToConfirm*). Notice that the *Impact* category is assigned by the person that records the incident in the incident management system at the time of its creation.

Due to the fact that those schedules are not normally distributed, we can only perform a non-parametric analysis of variance. We will use the Kruskal-Wallis one-way analysis of variance, an extension of the Mann-Whitney U test, which is the nonparametric analog of one-way ANOVA test. The Kruskal-Wallis H test allows assessing whether several independent samples are from the same population (i.e. if they have similar statistical distributions). In our case those independent samples are the groups of incidents for each of the four *Impact* categories.

Let T be a schedule and i and j two different impact categories. Then, the underlying hypotheses for this test are the following:

$$H_0:\ \forall_{i,j}:T_i \sim T_j \qquad vs. \qquad H_1:\ \neg\ \forall_{i,j}:T_i \sim T_j$$

Table 5. Testing the influence of the impact on incident schedules
with the Kruskal-Wallis one-way analysis of variance test

	TimeToRespond	TimeToResolve	TimeToConfirm
Chi-Square	352.381	77.532	18.487
df	3	3	3
Asymp. Sig.	.000	.000	.000

The Kruskal-Wallis H test statistic is distributed approximately as chi-square. Consulting a chi-square table with df = 3 (degrees of freedom) and for a significance of $\alpha = 0.01$ (probability of Type I error of 1%) we obtain a critical value of chi-square of

11.3. Since this value is less than the computed H values (for each of the schedule variables in Table 5), we reject the null hypothesis that the samples do not differ on the criterion variable (the *Impact*). In other words, given any of the schedule variables, we cannot sustain that the statistical distributions of the groups of incidents corresponding to each of the *Impact* categories are the same. This means that we accept the alternative hypothesis that **the impact of an incident has influence on all the schedule variables**.

RQ: Has the priority of an incident an influence on its lifecycle?
Here we want know if incidents with different assigned priorities differ in the corresponding lifecycle schedules (*TimeToRespond, TimeToResolve, TimeToConfirm*). We will follow the same rationale as for the previous research question, regarding the applicable statistic and its interpretation.

Table 6. Testing the influence of the priority on incident schedules
with the Kruskal-Wallis one-way analysis of variance test

	TimeToRespond	TimeToResolve	TimeToConfirm
Chi-Square	298.918	80.868	13.210
df	3	3	3
Asymp. Sig.	.000	.000	.004

Again the critical value of chi-square for (df = 3, α = 0.01) = 11.3. Since this value is less than the computed H values for each of the schedule variables in Table 6, we reject the null hypothesis that the samples do not differ on the criterion variable (the *Priority*). In other words, given any of the schedule variables, we cannot sustain that the statistical distributions of the groups of incidents corresponding to each of the *Priority* categories are the same. This means that we accept the alternative hypothesis that **the priority of an incident has influence on all the schedule variables**.

RQ: Has the originating country of an incident an influence on its lifecycle?
The rational for answering this research question is the same as for the previous one. To enable the application of the Kruskal-Wallis test, we have automatically recoded the *Country* variable from string categories into numerical categories.

Table 7. Testing the influence of the originating country on incident schedules
with the Kruskal-Wallis one-way analysis of variance test

	TimeToRespond	TimeToResolve	TimeToConfirm
Chi-Square	1666.912	337.181	44.877
df	5	5	5
Asymp. Sig.	.000	.000	.000

Given that the critical value of chi-square for (df = 5, α = 0.01) = 15.1. Since this value is less than the computed H values for each of the schedule variables in Table 7, we reject the null hypothesis that the samples do not differ on the criterion variable (the *Country*). In other words, given any of the schedule variables, we cannot sustain that the statistical distributions of the groups of incidents corresponding to each of the

countries are the same. This means that we accept the alternative hypothesis that **the country of an incident has influence on all the schedule variables**.

RQ: Has the originating geographical zone of an incident an influence on its lifecycle?
The rational for answering this research question is again the same as for the previous one. To enable the application of the Kruskal-Wallis test, we have automatically re-coded the *Zone* variable from string categories into numerical categories.

Table 8. Testing the influence of the originating zone on incident schedules
with the Kruskal-Wallis one-way analysis of variance test

	TimeToRespond	TimeToResolve	TimeToConfirm
Chi-Square	1546.415	139.297	17.727
df	2	2	2
Asymp. Sig.	.000	.000	.000

Given that the critical value of chi-square for (df = 2, α = 0.01) = 9.21, we reject the null hypothesis that the samples do not differ on the criterion variable (the *Zone*). In other words, given any of the schedule variables, we cannot sustain that the statistical distributions of the groups of incidents corresponding to each of the geographical zones are the same. Then we accept the alternative hypothesis that **the geographical zone where the incident was reported has influence on all the schedule variables**.

RQ: Has the incident category an influence on its lifecycle?
Again, after performing an automatic recode (for the *Category* variable), we obtained the following summary table:

Table 9. Testing the influence of the category on incident schedules
with the Kruskal-Wallis one-way analysis of variance test

	TimeToRespond	TimeToResolve	TimeToConfirm
Chi-Square	837.595	1258.178	612.215
df	15	15	15
Asymp. Sig.	.000	.000	.000

Given that the critical value of chi-square for (df = 15, α = 0.01) = 30.6, we reject the null hypothesis that the samples do not differ on the criterion variable (the incident *Category*). In other words, given any of the schedule variables, we cannot sustain that the statistical distributions of the groups of incidents corresponding to each category are the same. This means that we accept the alternative hypothesis that **the incident category has influence on all the schedule variables**.

RQ: Is the distribution of critical priority incidents the same across countries?
Since we know the proportion of total incident reports originated in each country (see Fig. 3) we can expect that the incidents with critical priority per country follow the same proportion of values. For this purpose we will use the Chi-Square Test proce-

dure that tabulates a variable into categories and computes a chi-square statistic. This non-parametric goodness-of-fit test compares the observed and expected frequencies in each country to test if each one contains the same proportion of values.
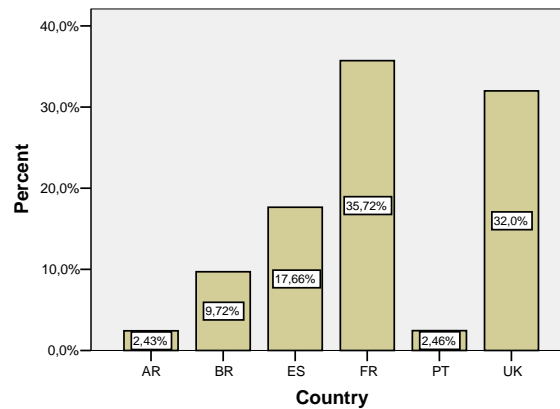

Fig. 3. Percentage of incident reports per country

To apply this test we only selected the critical incidents and obtained the results displayed in Table 10. Since the critical value of the chi-square for (df = 5, α = 0.01) = 15.1, we reject the null hypothesis that the proportion of critical priority incidents is the same across countries. This means that we accept the alternative hypothesis that **the proportion of critical priority incidents is different across countries**.

Table 10. Results of applying the Chi-Square Test procedure to assess if the distribution of critical priority incidents is the same across countries

	Observed N	Expected N	Residual
AR	12	17.8	-5.8
BR	39	71.2	-32.2
ES	154	129.3	24.7
FR	198	261.5	-63.5
PT	15	18.0	-3.0
UK	314	234.3	79.7
Total	732		

	Country
Chi-Square	64.203
df	5
Asymp. Sig.	.000

4 Conclusions and future work

4.1 Conclusions

In this paper we obtained statistically significant evidence that several independent variables (*Impact, Priority, Country, Zone* and *Category*) have an influence on incidents lifecycle, as characterized by three dependent variables (*TimeToRespond, Time-ToResolve* and *TimeToConfirm*). To assess the intensity of the relationship among the

independent and dependent variables we must use appropriate measures of association, but that analysis could not be included in this paper due to space restrictions.

There is no surprise on the influence of incident's business criticality (the *Impact*) and incident's correction prioritization recorded by the support (the *Priority*) on incidents lifecycle. After all, those incident descriptors were proposed with that same aim.

Not so obvious is the observed fact that either the country or the geographical zone of an organization reporting an incident, has influence on all descriptive variables that characterize incidents lifecycle. This means that organizations from different countries (or geographical zones) do not receive the same kind of support, although they are using the same products and, in principle, paying approximately the same for it. Several reasons, which we have not been explored yet, may explain this phenomenon:

- exigency on SLAs formalization and compliance verification by clients may somehow differ from country to country;
- cultural differences that cause a distinction on the tolerance to failure by final users (e.g. not complaining because an incident was yet solved);
- language differences that somehow influence the relationship between final users and the international support that is provided by the software vendor worldwide,

The incident category also has a direct influence on the three schedule variables. However, we have many kinds of recorded incidents, ranging from those occurring at software installation, to those related to software functionalities. The incidents can also go from enhancement requests to "true" bugs. This diversity requires a careful study before any interpretation of value can be performed.

Another apparent surprise was the fact that the proportion of critical incidents is not the same across countries. In all countries, except the UK and Spain, the actual number of critical incidents was below the expectation. This may indicate that end-users in those countries are causing an over-grading in incidents critically assessment by the support. Sometimes, end-users/customers tend to think that their incidents have always higher impact, simply because it affects the way they do their work and not based on the impact the incident has on the business. Again, this issue deserves further study before sensible conclusions can be drawn.

We have taken a view of the incident management process inspired by the ITIL approach, thus highlighting the importance of combining efforts to link engineering and management areas.

4.2 Threats to the validity

The main threats to this empirical study are related with data quality and the incident management process itself.

The main data quality related threats are:

- Data missing and/or wrong data (product name, version, etc) provided from the end-users/customers;
- Wrong data entered by the support staff (priority, impact, categorization, resolution codes, etc).

The main Incident Management process threats are:

- Lack of skills about the support tool can make some information non reliable (time to respond to incidents, time to resolve, etc);
- Customer non-response to a provided solution can cause incidents to be open when in fact they could be closed.

As an external threat to this empirical study, we can point that there is data missing from the software development process (resources allocated, activities, development tools, development methodology, etc.) which could help us to better evaluate and understand some of the results.

4.3   Future work

This empirical study was built upon a large sample of real-life data on incidents across a large period of time, on a long list of commercial products and customers in different countries. We are conscious that we have only scratched the surface. We plan to continue this work by deeply analyzing all the incidents, their categories, software errors and their causes.

Besides understanding the incident management process, our final aim is proposing some guidelines to cost-effectively improve software quality, based on incident management optimization. These guidelines can be focused on the products that appear to have more reported incidents or simply based on the most frequent incident categories. For this to be done accurately, we plan to collect more data, such as information about software development resources and activities performed during the overall development process.

References
1.   Humphrey, W.: Managing the Software Process. Addison-Wesley Publishing Company, Boston, MA (1989)
2.   El-Eman, K., Drouin, J.-N., and Melo, W. (eds.): SPICE: The Theory and Practice of Software Process Improvement and Capability Determination. IEEE Computer Society Press, Los Alamitos, CA (1997)
3.   Sjøberg, D.I.K., Hannay, J.E., Hansen, O., Kampenes, V.B., Karahasanovic, A., Liborg, N.-K., and Rekdal, A.: A survey of controlled experiments in software engineering. IEEE Transactions on Software Engineering, vol. 31, nr. 9, pp. 733-753 (2005)
4.   Jedlitschka, A., and Ciolkowski, M.: Towards Evidence in Software Engineering. In: Proc. of International Symposium on Empirical Software Engineering (ISESE'04), Washington, DC, pp. 261-270. IEEE Computer Society (2004)
5.   OGC: The Official Introduction to the ITIL Service Lifecycle Book. TSO, London (2007)
6.   Cannon, D., Wheeldon, D: ITIL Service Operation. TSO, London (2007)
7.   Case, G., Spalding, G.: ITIL Continual Service Improvement. TSO, London (2007)
8.   Iqbal, M., Nieves, M.: ITIL Service Strategy. TSO, London (2007)

9. Lacy, S., MacFarlane I.: ITIL Service Transition. TSO, London (2007)
10. Loyd, V., Ruud, C.: ITIL Service Design. TSO, London (2007)
11. Abran, A., Moore, J.W., Bourque, P., and Dupuis, R. (eds.): Guide to the Software Engineering Body of Knowledge (SWEBOK). IEEE Computer Society (2004)
12. Abe, T.: What is Service Science? Research report nr. 246, Fujitsu Research Institute, Tokyo (2005)
13. Poole, G.: A new academic discipline needed for the 21st century. Triangle Business Journal (2007)
14. Barash, G., Bartolini, C., and Wu, L.: Measuring and Improving the Performance of an IT Support Organization in Managing Service Incidents. In: Proc. of 2nd IEEE/IFIP International Workshop on Business-Driven IT Management (BDIM '07), pp. 11-18. IEEE Computer Society (2007)
15. Niessink, F., and Vliet, H.v.: Software Maintenance from a Service Perspective. Journal of Software Maintenance: Research and Practice, vol. 12, nr. 2, pp. 103-120. Wiley InterScience (2000)
16. Jansen, S., and Brinkkemper, S.: Evaluating the Release, Delivery and Development Processes of Eight Large Product Software Vendors applying the Customer Configuration Update Model. In: Proc. of International Workshop on Interdisciplinary Software Engineering Research (WISER'06) @ ICSE'06, Shanghai, China (2006)
17. Mohagheghi, P., and Conradi, R.: Quality, productivity and economic benefits of software reuse: a review of industrial studies. Empirical Software Engineering, vol. 12, nr. 5, pp. 471-516. Kluwer Academic Publishers, Hingham, MA (2007)